

Orbital optimization in electronic wave functions; equations for quadratic and cubic convergence of general multiconfiguration wave functions*

Laurence G. Yaffe¹ and William A. Goddard III

Arthur Amos Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, California 91125

(Received 3 November 1975)

We derive variational equations for optimization of the orbitals of arbitrary multiconfiguration wave functions. Expressing the transformation matrix connecting the set of orthonormal trial vectors and the set of final optimal orbitals as an exponential matrix of independent rotation angles allows a simple derivation of the coupled variational equations to arbitrary order. We include the explicit results through third order which are sufficient for cubic convergence in the iterative solution of the optimum orbitals. The equations were programmed (including all terms), and applications to Hartree-Fock and generalized valence bond wave functions of the carbon atom, FeO₂, and NiCH₂ are reported.

I. INTRODUCTION

A number of standard methods for obtaining electronic wave functions involve the optimization of orthogonal orbitals for some given form of the wave function. Examples include Hartree Fock (HF), generalized valence bond (GVB), and multi-configuration-self-consistent-field (MC-SCF) wave functions.

Any electronic wave function constructed from orthonormal orbitals $\{\chi_\alpha\}$ leads to an energy expression of the form

$$E = \sum_{\alpha, \beta} \left(2f_{\alpha\beta} \langle \alpha | h | \beta \rangle + \sum_{\gamma, \delta} c_{\alpha\beta\gamma\delta} \langle \alpha\beta | \gamma\delta \rangle \right), \quad (1)$$

where the orbitals are denoted by their subscripts, and $\langle \alpha | h | \beta \rangle$ and $\langle \alpha\beta | \gamma\delta \rangle$ are the one- and two-electron components of the energy

$$\langle \alpha | h | \beta \rangle = \int d\tau_1 \chi_\alpha(1) h(1) \chi_\beta(1),$$

$$\langle \alpha\beta | \gamma\delta \rangle = \int d\tau_1 d\tau_2 \chi_\alpha(1) \chi_\beta(1) (1/r_{12}) \chi_\gamma(2) \chi_\delta(2)$$

(for convenience we take the orbitals to be real).

For specific types of wave functions, including general open-shell HF and a restricted class of GVB wave functions, a much simpler energy expression is obtained¹

$$E = \sum_{\alpha} 2f_{\alpha} \langle \alpha | h | \alpha \rangle + \sum_{\alpha, \beta} (a_{\alpha\beta} J_{\alpha\beta} + b_{\alpha\beta} K_{\alpha\beta}), \quad (2)$$

where $J_{\alpha\beta} = \langle \alpha\alpha | \beta\beta \rangle$ and $K_{\alpha\beta} = \langle \alpha\beta | \alpha\beta \rangle$ are the usual Coulomb and exchange integrals.

The simple closed-shell HF wave function (all orbitals doubly occupied, once with spin up and once with spin down) leads to

$$E = \sum_{\alpha} 2\langle \alpha | h | \alpha \rangle + \sum_{\alpha, \beta} (2J_{\alpha\beta} - K_{\alpha\beta}). \quad (2')$$

In order to optimize the wave function, one starts with a set of trial functions $\{\chi_\alpha\}$ and attempts to find a transformation T which produces a new set of orbitals $\{\varphi_i\}$:

$$\varphi_i = \sum_{\alpha} T_{i\alpha} \chi_\alpha \quad (3)$$

such that the resulting energy is stationary.

A common approach,² often used to solve for the optimum wave function, begins by varying each orbital independently (keeping all other occupied orbitals fixed) and finding the resulting change in energy. This is equivalent to taking the derivative of the energy with respect to each component of the transformation matrix T and adding constraints to ensure orthogonality to the other occupied orbitals. Upon requiring the energy to be stationary under such variations, one obtains a set of pseudoeigenvalue equations of the form^{2,3}

$$\hat{H}^i \varphi_i = \sum_j \epsilon_{ij} \varphi_j, \quad \text{all } i \quad (4)$$

(where $\epsilon_{ij} = \epsilon_{ji}$ and where \hat{H}^i depends on all the orbitals), which are satisfied by the optimum orbitals $\{\varphi_i\}$. In order to solve (4), the usual approach is to evaluate the operators \hat{H}^i using the trial functions $\{\chi_\alpha\}$, then solve for the resulting orbitals, and continue the process until convergence is achieved (hopefully). This approach has been particularly successful for the HF wave function (2') [in that case all the one-electron operators (\hat{H}^i) are identical]. However, Eq. (4) in no way specifies the best way in which to modify a set of nonoptimum orbitals. Thus, even a good starting guess may converge quite slowly or not at all.

In order to obtain variational equations which actually specify how to correct a set of trial functions, one must expand the energy beyond first or-

der in the orbital variations. However, to obtain sufficient conditions one must allow each orbital to vary in *all* directions, including those occupied by other orbitals. Because of the orthogonality constraints, such variations require simultaneous changes in more than one orbital.

Hunt, Goddard, and Dunning⁴ constructed the transformation matrix as $\underline{T} = \underline{1} + \underline{\Lambda}$ (where $\underline{\Lambda}$ is antisymmetric, thus incorporating orthogonality through first order into the transformation matrix) and obtained variational equations specifying how to simultaneously correct all orbitals (through first order). These equations would lead to quadratic convergence but were not implemented. By using a restricted form of these results, valid for pairwise (two by two) rotations, Hunt, Hay, and Goddard¹ developed a general program (GVBONE) capable of optimizing open-shell HF and restricted (perfect pairing) GVB wave functions [i.e., wave functions with energies expressible as (2)]. However, owing to neglect of coupling between pairwise rotations, the convergence is not quadratic.

Bobrowicz and Goddard^{5,6} used an alternative approach of expanding the energy through second order in all orbital corrections and incorporating all orthonormality conditions explicitly, leading also to equations sufficient for quadratic convergence. Using approximate forms of these equations, Bobrowicz, Wadt, and Goddard⁶ developed a general program (GVB TWO) capable of more rapid convergence than GVBONE but still far from quadratic.

In the above approaches the transformation matrix components themselves are varied directly. Since these components are not independent (owing to the required unitarity of \underline{T}), it is very difficult to obtain consistent variational equations correct to an arbitrary order.

Our approach begins with expressing the transformation matrix \underline{T} in terms of completely independent rotation angles Δ_{ij} in such a way that orbital orthogonality is completely incorporated into the form of the transformation matrix. By varying the independent rotation angles Δ_{ij} instead of the components of \underline{T} consistent variational equations (valid to any given order) may then be derived for either the general (1) or the restricted (2) energy expression. Terms up to third order in the energy are explicitly expanded, thus giving cubically convergent variational equations. These equations were programmed (including all terms) and the results of several applications are discussed.

II. ROTATION MATRICES

Given the form of a wave function expressed in terms of N orthonormal orbitals, we wish to determine the optimum orthonormal orbitals $\{\varphi_i\}$.

Beginning with a set of N orthonormal trial functions $\{\chi_\alpha\}$, the relationship of the optimum orbitals $\{\varphi_i\}$ to $\{\chi_\alpha\}$ can be expressed in terms of a transformation matrix \underline{T}

$$\varphi_i = T_{i\alpha} \chi_\alpha \quad (5)$$

[in (5) and subsequently we will generally use the summation convention for repeated indices; for emphasis we will occasionally indicate the summation explicitly].

Since both sets of orbitals are orthonormal, \underline{T} is a unitary transformation. Because the sign of an orbital is of no consequence, we need only consider special unitary transformations $SU(N)$. Since we take our orbitals to be real, only orthogonal transformations in $SO(N)$ need be considered.

An arbitrary element of $SO(N)$ may be constructed from $N(N-1)/2$ independent generators. In order to derive unconstrained variational equations, we need to express \underline{T} directly in terms of $N(N-1)/2$ independent parameters specifying the transformation.

We may take the generators to be the set of all pairwise infinitesimal rotations of the N basis vectors for the N -dimensional space (i.e., each generator is an infinitesimal rotation in the plane containing two basis vectors). The transformation matrix corresponding to an infinitesimal rotation of basis vectors φ_α and φ_β ($\alpha > \beta$) is given by

$$\underline{T}^{\alpha\beta} = \underline{1} + \underline{\Lambda}^{\alpha\beta} d\theta, \quad (6)$$

where

$$\Lambda_{ij}^{\alpha\beta} = \delta_{i\alpha} \delta_{j\beta} - \delta_{i\beta} \delta_{j\alpha} \quad (7)$$

(that is, all but two of the N^2 elements of $\underline{\Lambda}^{\alpha\beta}$ are zero, the nonzero components being $\Lambda_{\alpha\beta}^{\alpha\beta} = +1$, and $\Lambda_{\beta\alpha}^{\alpha\beta} = -1$).

An arbitrary infinitesimal transformation $d\hat{R}$ may be formed by a linear combination of the generators. Thus the transformation matrix is given by

$$\underline{T}^{d\hat{R}} = \underline{1} + \underline{\Lambda}^{\hat{R}} d\theta, \quad (8)$$

where

$$\underline{\Lambda}^{\hat{R}} = \sum_{\alpha > \beta} \gamma_{\alpha\beta} \underline{\Lambda}^{\alpha\beta} \quad (9)$$

and

$$\sum_{\alpha > \beta} (\gamma_{\alpha\beta})^2 = 1. \quad (10)$$

The representation of a finite transformation \hat{R} is constructed by repeated application of the infinitesimal transformation $d\hat{R}$. Thus

$$\begin{aligned} \underline{T}^{\hat{R}} &= \lim_{n \rightarrow \infty} [1 + \underline{\Lambda}^{\hat{R}} \theta / n]^n \\ &= 1 + (\underline{\Lambda}^{\hat{R}} \theta) + \frac{1}{2} (\underline{\Lambda}^{\hat{R}} \theta)^2 + \frac{1}{6} (\underline{\Lambda}^{\hat{R}} \theta)^3 + \dots \end{aligned} \quad (11)$$

Hence,

$$\underline{T}^{\hat{R}} = \exp(\underline{\Lambda}^{\hat{R}} \theta). \quad (12)$$

However, $\underline{\Lambda}^{\hat{R}} \theta$ simply specifies an arbitrary antisymmetric matrix. Thus a general form for an N -dimensional proper orthogonal matrix is given by^{7,8}

$$\underline{T} = e^{\underline{\Delta}}, \quad (13)$$

where $\underline{\Delta}$ is any antisymmetric matrix.

For convenience of discussion such a transformation will be referred to as a rotation, and the $N(N-1)/2$ independent elements of $\underline{\Delta}$ (all $\Delta_{\alpha\beta}$ with $\alpha > \beta$) referred to as rotation angles. However, such a transformation need not be an actual N -dimensional rotation occurring in some N -dimensional plane.

III. ORBITAL OPTIMIZATION

A. General energy expression

The energy (1) of a wave function constructed from orthogonal orbitals may be put in the form

$$E = \sum_{\alpha, \beta} (f_{\alpha\beta} h_{\alpha\beta} + H_{\alpha\beta}^{\alpha\beta}), \quad (14)$$

where

$$\begin{aligned} h_{\alpha\beta} &= \langle \alpha | h | \beta \rangle, \\ H_{\mu\nu}^{\alpha\beta} &= f_{\alpha\beta} \langle \mu | h | \nu \rangle + \sum_{\gamma, \delta} c_{\alpha\beta\gamma\delta} (\mu\nu | \gamma\delta), \end{aligned} \quad (15)$$

and where the functions $\{\chi_\alpha\}$ are denoted by their subscripts only.

Applying a transformation \underline{T} to the orbitals $\{\chi_\alpha\}$

$$\varphi_i = \sum_{\alpha} T_{i\alpha} \chi_\alpha \quad (16)$$

leads to a new energy

$$\begin{aligned} E' &= \sum_{i,j} \left(2 f_{ij} \langle \varphi_i | h | \varphi_j \rangle + \sum_{k,l} c_{ijkl} (\varphi_i \varphi_j | \varphi_k \varphi_l) \right) \\ &= \sum_{i,j} T_{i\alpha} T_{j\beta} \left(2 f_{ij} h_{\alpha\beta} + \sum_{k,l} T_{k\gamma} T_{l\delta} c_{ijkl} (\alpha\beta | \gamma\delta) \right). \end{aligned} \quad (17)$$

In order for the energy to be stationary, we require⁹

$$\begin{aligned} 0 &= \frac{\partial E}{\partial \Delta_{\mu\nu}} \\ &= 4 \sum_{i,j} \frac{\partial T_{i\alpha}}{\partial \Delta_{\mu\nu}} T_{j\beta} \left(f_{ij} h_{\alpha\beta} + \sum_{k,l} T_{k\gamma} T_{l\delta} c_{ijkl} (\alpha\beta | \gamma\delta) \right) \end{aligned}$$

or

$$0 = \frac{\partial T_{i\alpha}}{\partial \Delta_{\mu\nu}} T_{j\beta} H_{\alpha\beta}^{i\sigma j}, \quad (18)$$

for all $\mu > \nu$, where

$$H_{\alpha\beta}^{i\sigma j} = f_{ij} h_{\alpha\beta} + c_{ijkl} (\alpha\beta | \varphi_k \varphi_l) \quad (19)$$

and summations over i, j, k , and l are implicit.

In order to solve (18) we must express the equations in terms of the *independent* rotation angles (all Δ_{ij} with $i > j$). Thus expanding \underline{T} leads to

$$\begin{aligned} T_{j\beta} &= \delta_{j\beta} + \Delta_{j\beta} + \frac{1}{2} \Delta_{jk} \Delta_{k\beta} + \frac{1}{6} \Delta_{jk} \Delta_{kl} \Delta_{l\beta} + O(\Delta^4) \\ &= \delta_{j\beta} + \Delta_{mn} P_{mn}^{j\beta} + \frac{1}{2} \Delta_{mn} \Delta_{pq} Q_{mnpq}^{j\beta} \\ &\quad + \frac{1}{6} \Delta_{mn} \Delta_{pq} \Delta_{rs} S_{mnpqrs}^{j\beta} + O(\Delta^4), \end{aligned} \quad (20)$$

where

$$m > n, \quad p > q, \quad r > s;$$

$$P_{mn}^{j\beta} = (1 - \tau_{mn})(\delta_{jm} \delta_{n\beta}), \quad (21)$$

$$Q_{mnpq}^{j\beta} = \frac{1}{2} (1 + \tau_{mp} \tau_{nq})(1 - \tau_{mn})(1 - \tau_{pq})(\delta_{jm} \delta_{n\beta} \delta_{qp}), \quad (22)$$

$$\begin{aligned} S_{mnpqrs}^{j\beta} &= \frac{1}{12} (1 + \tau_{mp} \tau_{nq} + \tau_{mr} \tau_{ns} + \tau_{pr} \tau_{qs} \\ &\quad + \tau_{mr} \tau_{ns} \tau_{mp} \tau_{nq} + \tau_{pr} \tau_{qs} \tau_{mp} \tau_{nq}) \\ &\quad \times (1 - \tau_{mn})(1 - \tau_{pq})(1 - \tau_{rs})(\delta_{jm} \delta_{n\beta} \delta_{qr} \delta_{s\beta}) \end{aligned} \quad (23)$$

($\tau_{\alpha\beta}$ is the permutation operator interchanging α and β). Similarly,

$$\frac{\partial T_{i\alpha}}{\partial \Delta_{\mu\nu}} = P_{\mu\nu}^{i\alpha} + \Delta_{mn} Q_{\mu\nu mn}^{i\alpha} + \Delta_{mn} \Delta_{pq} S_{\mu\nu mn pq}^{i\alpha} + O(\Delta^3), \quad (24)$$

$$\begin{aligned} H_{\alpha\beta}^{i\sigma j} &= H_{\alpha\beta}^{ij} + 2 \Delta_{mn} P_{mn}^{i\sigma} V_{\alpha\beta k\gamma}^{ijkl} \\ &\quad + \Delta_{mn} \Delta_{pq} (P_{mn}^{k\sigma} P_{pq}^{l\delta} V_{\alpha\beta\gamma\delta}^{ijkl} + Q_{mnpq}^{i\sigma} V_{\alpha\beta k\gamma}^{ijkl}) + O(\Delta^3), \end{aligned} \quad (25)$$

where

$$V_{\alpha\beta\gamma\delta}^{ijkl} = c_{ijkl} (\alpha\beta | \gamma\delta) \quad (26)$$

Substituting (24), (20), and (25) into (18) leads to

$$\begin{aligned}
0 &= [P_{\mu\nu}^{i\alpha} + \Delta_{mn} Q_{\mu\nu mn}^{i\alpha} + \Delta_{mn} \Delta_{pq} S_{\mu\nu mn pq}^{i\alpha}] [\delta_{i\beta} + \Delta_{mn} P_{mn}^{j\beta} + \frac{1}{2} \Delta_{mn} \Delta_{pq} Q_{mn pq}^{j\beta}] \\
&\quad \times [H_{\alpha\beta}^{ij} + 2\Delta_{mn} P_{mn}^{i\gamma} V_{\alpha\beta k\gamma}^{ijk} + \Delta_{mn} \Delta_{pq} (P_{mn}^{k\gamma} P_{pq}^{i\delta} + \delta_{i\delta} Q_{mn pq}^{k\gamma}) V_{\alpha\beta\gamma\delta}^{ijkl}] + O(\Delta^3), \\
0 &= [P_{\mu\nu}^{i\alpha} H_{\alpha j}^{ij}] + \Delta_{mn} [Q_{\mu\nu mn}^{i\alpha} H_{\alpha j}^{ij} + P_{\mu\nu}^{i\alpha} P_{mn}^{j\beta} (H_{\alpha\beta}^{ij} + 2V_{\alpha k i j}^{ijk})] \\
&\quad + \Delta_{mn} \Delta_{pq} [S_{\mu\nu mn pq}^{i\alpha} H_{\alpha j}^{ij} + \frac{1}{2} (Q_{\mu\nu mn}^{i\alpha} P_{pq}^{j\beta} + Q_{\mu\nu pq}^{i\alpha} P_{mn}^{j\beta} + Q_{mn pq}^{i\alpha} P_{\mu\nu}^{j\beta}) (H_{\alpha\beta}^{ij} + 2V_{\alpha k i j}^{ijk})] \\
&\quad + (P_{\mu\nu}^{i\alpha} P_{mn}^{j\beta} P_{pq}^{k\gamma} + P_{\mu\nu}^{i\alpha} P_{pq}^{j\beta} P_{mn}^{k\gamma} + P_{mn}^{i\alpha} P_{pq}^{j\beta} P_{\mu\nu}^{k\gamma}) V_{\alpha\beta\gamma\delta}^{ijkl} + O(\Delta^3). \tag{27}
\end{aligned}$$

Consequently

$$0 = X_{\mu\nu} + \Delta_{mn} B_{\mu\nu mn} + \Delta_{mn} \Delta_{pq} C_{\mu\nu mn pq} + O(\Delta^3), \tag{28a}$$

where

$$\mu > \nu, \quad m > n, \quad p > q,$$

$$X_{\mu\nu} = (1 - \tau_{\mu\nu}) H_{\nu j}^{\mu j}, \tag{28b}$$

$$B_{\mu\nu mn} = \frac{1}{2} (1 + \tau_{\mu m} \tau_{\nu n}) (1 - \tau_{\mu\nu}) (1 - \tau_{mn}) [H_{\nu n}^{\mu m} + 2V_{\nu j k n}^{\mu j k m} + \delta_{\mu\nu} H_{n j}^{\mu j}], \tag{28c}$$

$$C_{\mu\nu mn pq} = \frac{1}{2} (1 + \tau_{\mu m} \tau_{\nu n} + \tau_{\mu p} \tau_{\nu q}) (1 + \tau_{m p} \tau_{n q}) (1 - \tau_{\mu\nu}) (1 - \tau_{mn}) (1 - \tau_{pq}) [V_{\nu n q j}^{\mu m p j} + \delta_{\mu\nu} (\frac{1}{2} H_{n q}^{\mu p} + V_{n j k q}^{\mu j k p} + \frac{1}{6} \delta_{m n} H_{q j}^{\mu j})]. \tag{28d}$$

Note that X , B , and C are antisymmetric under interchange of pair indices (i.e., $X_{\mu\nu} = -X_{\nu\mu}$, $B_{\mu\nu mn} = -B_{\mu\nu nm}$, etc.) but are symmetric under interchange of pairs of indices (i.e., $B_{\mu\nu mn} = B_{mn\mu\nu}$, $C_{\mu\nu mn pq} = C_{mn pq\mu\nu}$, etc.).¹⁰ Applying the permutation operators in (28) leads to the expanded form of the variational equations contained in Appendix (A1).

B. Restricted energy expression

The restricted energy expression (2) may be put in the form

$$E = \sum_{\alpha} (f_{\alpha} h_{\alpha\alpha} + H_{\alpha\alpha}^{\alpha}), \tag{29}$$

where

$$H_{\mu\nu}^{\alpha} = f_{\alpha} h_{\mu\nu} + \sum_{\gamma} (a_{\alpha\gamma} J_{\mu\nu}^{\gamma} + b_{\alpha\gamma} K_{\mu\nu}^{\gamma}). \tag{30}$$

Thus, for the energy to be stationary we require that

$$0 = \frac{\partial E}{\partial \Delta_{\mu\nu}} = 4 \frac{\partial T_{i\alpha}}{\partial \Delta_{\mu\nu}} T_{i\beta} H_{\alpha\beta}^{\mu}, \quad \mu > \nu \tag{31}$$

where

$$\begin{aligned}
H_{\alpha\beta}^{\mu} &= f_{\mu} h_{\alpha\beta} + \sum_j [a_{i j} (\alpha\beta | \varphi_j \varphi_j) + b_{i j} (\alpha\varphi_j | \beta\varphi_j)] \\
&= H_{\alpha\beta}^i + 2\Delta_{mn} P_{mn}^{i\gamma} g_{\alpha\beta j\gamma}^{ij} + \Delta_{mn} \Delta_{pq} (P_{mn}^{i\gamma} P_{pq}^{j\delta} g_{\alpha\beta\gamma\delta}^{ij} + Q_{mn pq}^{i\gamma} g_{\alpha\beta j\gamma}^{ij}) + O(\Delta^3)
\end{aligned} \tag{32}$$

and

$$g_{\alpha\beta\gamma\delta}^{ij} = a_{i j} (\alpha\beta | \gamma\delta) + \frac{1}{2} b_{i j} [(\alpha\gamma | \beta\delta) + (\alpha\delta | \beta\gamma)]. \tag{33}$$

Therefore,

$$\begin{aligned}
0 &= [P_{\mu\nu}^{i\alpha} H_{\alpha i}^i] + \Delta_{mn} [Q_{\mu\nu mn}^{i\alpha} H_{\alpha i}^i + P_{\mu\nu}^{i\alpha} P_{mn}^{j\beta} (\delta_{ij} H_{\alpha\beta}^i + 2g_{\alpha i j\beta}^{ij})] \\
&\quad + \Delta_{mn} \Delta_{pq} [S_{\mu\nu mn pq}^{i\alpha} H_{\alpha i}^i + \frac{1}{2} (Q_{\mu\nu mn}^{i\alpha} P_{pq}^{j\beta} + Q_{\mu\nu pq}^{i\alpha} P_{mn}^{j\beta} + Q_{mn pq}^{i\alpha} P_{\mu\nu}^{j\beta}) (\delta_{ij} H_{\alpha\beta}^i + 2g_{\alpha i j\beta}^{ij})] \\
&\quad + (P_{\mu\nu}^{i\alpha} P_{mn}^{j\beta} P_{pq}^{i\gamma} + P_{\mu\nu}^{i\alpha} P_{pq}^{j\beta} P_{mn}^{i\gamma} + P_{mn}^{i\alpha} P_{pq}^{j\beta} P_{\mu\nu}^{i\gamma}) g_{\alpha i \beta\gamma}^{ij} + O(\Delta^3). \tag{34}
\end{aligned}$$

Thus¹¹

$$0 = X_{\mu\nu} + \Delta_{mn} B_{\mu\nu mn} + \Delta_{mn} \Delta_{pq} C_{\mu\nu mn pq} + O(\Delta^3), \quad (35a)$$

where $\mu > \nu$, $m > n$, $p > q$,

$$X_{\mu\nu} = (1 - \tau_{\mu\nu}) H_{\mu\nu}^{\mu}, \quad (35b)$$

$$B_{\mu\nu mn} = \frac{1}{2}(1 + \tau_{\mu m} \tau_{\nu n})(1 - \tau_{\mu\nu})(1 - \tau_{mn}) \\ \times [\delta_{\mu m}(H_{\nu n}^{\mu} - H_{\nu n}^{\nu} - H_{\nu n}^{\nu}) + 2G_{\mu\nu mn}^{\mu}], \quad (35c)$$

$$C_{\mu\nu mn pq} = \frac{1}{2}(1 + \tau_{\mu m} \tau_{\nu n} + \tau_{\mu p} \tau_{\nu q})(1 + \tau_{m p} \tau_{n q}) \\ \times (1 - \tau_{\mu\nu})(1 - \tau_{mn})(1 - \tau_{pq}) \\ \times \delta_{\mu m} [(G_{pq\nu n}^{\mu} - G_{pq\nu n}^{\nu}) + \frac{1}{2}\delta_{\nu n}(H_{\mu m}^{\mu} - \frac{1}{3}H_{\mu m}^{\nu})]. \quad (35d)$$

Applying the permutation operators in (35) leads to the expanded form of the variational equations in Appendix (A 2).

C. Pairwise rotations

If only a single rotation (say between orbitals 1 and 2) is to be performed, then the general energy expression (14) reduces to

$$E = E_0 + (f_{11}\bar{h}_{11} + H_{11}^{11}) + (f_{22}\bar{h}_{22} + H_{22}^{22}) + 2(f_{12}\bar{h}_{12} + H_{12}^{12}), \quad (36)$$

where E_0 contains all terms independent of orbitals 1 and 2. Optimizing the rotation between orbitals 1 and 2, one finds

$$0 = X_{21} + (\Delta_{21})B_{2121} + (\Delta_{21})^2 C_{212121}, \quad (37a)$$

where

$$X_{21} = H_{21}^{22} - H_{21}^{11} + H_{21}^{21} - H_{22}^{21}, \quad (37b)$$

Since

$$X_{\mu\nu} + B_{\mu\nu mn} \Delta_{mn} + C_{\mu\nu mn pq} \Delta_{mn} \Delta_{pq} = 0,$$

the contribution of rotation $(\mu\nu)$ to the change in the energy becomes

$$\Delta E^{\mu\nu} = -(\Delta_{\mu\nu})^2 (\frac{1}{2}B_{\mu\nu \mu\nu} + \Delta_{mn} C_{\mu\nu \mu\nu mn}) \\ + (\Delta_{\mu\nu})^3 (\frac{1}{3}C_{\mu\nu \mu\nu \mu\nu}) + O(\Delta^4), \quad (40)$$

where μ and ν are fixed. Naturally

$$\Delta E^{\text{tot}} \neq \sum_{\mu\nu} \Delta E^{\mu\nu}. \quad (41)$$

$$B_{2121} = H_{22}^{11} + H_{22}^{22} - H_{11}^{11} - H_{22}^{22} - 4H_{12}^{12} \\ + 2c_{2121}(J_{11} - 2J_{12} + J_{22}) \\ + 4(c_{2221} - c_{2111})(J_{12}^1 - J_{12}^2) \\ + 2(c_{1111} + c_{2222} - 2c_{2211})K_{12}, \quad (37c) \\ C_{212121} = 2(H_{21}^{11} - H_{21}^{22} + H_{22}^{21} - H_{11}^{21}) \\ + 3(c_{2222} + c_{1111} - 2c_{2211} - 4c_{2121})(J_{12}^1 - J_{12}^2) \\ + 3(c_{1112} - c_{2221})(2J_{12} + 4K_{12} - J_{11} - J_{22}). \quad (37d)$$

If the restricted energy expression is used, then^{11,12a}

$$X_{21} = H_{21}^{22} - H_{21}^{11}, \quad (37e)$$

$$B_{2121} = H_{22}^{11} - H_{11}^{11} + H_{22}^{22} - H_{22}^{22} + 2(a_{11} + a_{22} - 2a_{12})K_{12} \\ + (b_{11} + b_{22} - 2b_{12})(J_{12} + K_{12}), \quad (37f)$$

$$C_{212121} = 2(H_{21}^{11} - H_{21}^{22}) \\ + 3(a_{11} + a_{22} - 2a_{12} + b_{11} + b_{22} - 2b_{12})(J_{12}^1 - J_{12}^2). \quad (37g)$$

D. Energy contributions

In the case of multiple simultaneous rotations, it is often useful in the analysis of the iterative process to have an idea of the contribution of each rotation to the total decrease in the energy.

Given a set of rotation angles $\{\Delta_{mn}\}$, the total energy is of the form

$$E = E^0 + \Delta_{mn} X_{mn} + \frac{1}{2} \Delta_{mn} \Delta_{pq} B_{mn pq} \\ + \frac{1}{3} \Delta_{mn} \Delta_{pq} \Delta_{rs} C_{mn pq rs} + O(\Delta^4) \quad (38)$$

where $m > n$, $p > q$, $r > s$.

Now suppose that all rotations *except* the $(\mu\nu)$ rotation were performed. The difference in energy would be

$$\Delta E^{\mu\nu} = \Delta_{\mu\nu} X_{\mu\nu} + \Delta_{\mu\nu} (\Delta_{mn} B_{\mu\nu mn} - \frac{1}{2} \Delta_{\mu\nu} B_{\mu\nu \mu\nu}) + \Delta_{\mu\nu} [\Delta_{mn} \Delta_{pq} C_{\mu\nu mn pq} - \Delta_{\mu\nu} (\Delta_{mn} C_{\mu\nu \mu\nu mn} - \frac{1}{3} \Delta_{\mu\nu} C_{\mu\nu \mu\nu \mu\nu})] + O(\Delta^4) \\ = \Delta_{\mu\nu} (X_{\mu\nu} + \Delta_{mn} B_{\mu\nu mn} + \Delta_{mn} \Delta_{pq} C_{\mu\nu mn pq}) - (\Delta_{\mu\nu})^2 (\frac{1}{2} B_{\mu\nu \mu\nu} + \Delta_{mn} C_{\mu\nu \mu\nu mn}) + (\Delta_{\mu\nu})^3 (\frac{1}{3} C_{\mu\nu \mu\nu \mu\nu}) + O(\Delta^4). \quad (39)$$

However, the relative sizes of the $\Delta E^{\mu\nu}$ may provide useful indications of the relative importance of various rotations.

IV. DISCUSSION

A. Order of the convergence

Repeated application of Eqs. (28) [or (35)] and (13) to a set of trial orbitals will cause the initial wave function to converge cubically toward the optimum wave function. [That is, sufficiently close to convergence the gradient of the energy

(\bar{X}) and the change in the orbitals ($\bar{\Delta}$) both decrease cubically from one iteration to the next.¹³ If the tensor \underline{C} is ignored, then the resulting convergence should be quadratic. This cubic and quadratic convergence has been demonstrated for several prototype systems as indicated in Tables I-III for calculations on C, FeO₂, and NiCH₂.

B. Techniques

Several techniques were used to solve the variational equations. The first of these involves forming the matrix \underline{B} and the vector \bar{X} , inverting the matrix \underline{B} and obtaining the first-order solution vector $\bar{\Delta}_1$, satisfying $0 = \bar{X} + \underline{B} \cdot \bar{\Delta}_1$. The matrix $\underline{C} \cdot \bar{\Delta}_1$ is then constructed and $(\underline{B} + \underline{C} \cdot \bar{\Delta}_1)$ inverted in order to obtain the second-order solution vector $\bar{\Delta}_2$, which satisfies $0 = \bar{X} + (\underline{B} + \underline{C} \cdot \bar{\Delta}_1) \cdot \bar{\Delta}_2$. Continuing the process is unnecessary since $\bar{\Delta}_2$ is already correct to order Δ^3 . Finally, the transformation matrix $\underline{T} = e^{\Delta_2}$ is constructed and applied in order to obtain the new orbitals $\{\phi_i\}$.

Alternatively, one may solve the system of equations iteratively by minimizing the scalar functional

$$I[\bar{\Delta}] = \bar{X} \cdot \bar{\Delta} + \frac{1}{2} \underline{B} \cdot \bar{\Delta} \cdot \bar{\Delta} + \frac{1}{3} \underline{C} \cdot \bar{\Delta} \cdot \bar{\Delta} \cdot \bar{\Delta}, \quad (42)$$

where I is the total energy lowering. With this approach, each component of $\bar{\Delta}$ is optimized individually, and the process continued until the changes in each component are sufficiently small.

Both approaches were found to work satisfactorily. On the other hand, minimizing the functional $I = (\bar{X} + \underline{B} \cdot \bar{\Delta} + \underline{C} \cdot \bar{\Delta} \cdot \bar{\Delta})^2$, where I is the square of the error vector, was found to converge more slowly.

C. Results

The above techniques were applied to calculations on C, FeO₂, and NiCH₂ with the results shown in Tables I-III. "Hamiltonian diagonalization" refers to the usual optimization technique.¹ For "quadratic" and "cubic" calculations, the matrix \underline{B} and the tensors \underline{B} and \underline{C} , respectively, were evaluated. "Uncoupled" implies that only the diagonal elements of the matrix \underline{B} were evaluated (and used).^{12b}

For carbon atom we used a double zeta basis¹⁴ (ten basis functions) and solved for the HF wave function of the ³D state. Here the trial function was the GVB(1) wave function¹⁵ of C(³P), a good starting guess.

For FeO₂ we employed an effective potential¹⁶ to replace the Fe core electrons (reducing the system to 24 electrons) and used a double zeta basis set (33 basis functions). We carried out a GVB(3) calculation¹⁵ for the lowest singlet state

TABLE I. Convergence of HF wave function for the ¹D state of carbon atom. The starting vectors were obtained from the GVB(1) wave function for the ³P state. Add -37.0 to the energy quoted to obtain the total energy (in Hartree). \bar{X}^2 is the square of the gradient of the energy ($\bar{X}^2=0$ is the necessary condition for an optimum wave function). $\bar{\Delta}^2$ is the square of the change in the wave function.

Iteration	1	2	3	4	5	6	7	8	9
A. Hamiltonian diagonalization	Energy	-0.626 048	-0.626 944	-0.627 015	-0.627 017	-0.627 017	-0.627 017	-0.627 017	-0.627 017
	$\bar{\Delta}^2$	1.7E-3	6.2E-5	1.4E-6	2.8E-8	5.1E-10	9.1E-12	2.3E-13	9.1E-15
B. Quadratic	Energy	-0.626 226	-0.627 017	-0.627 017	-0.627 017	-0.627 017	-0.627 017	-0.627 017	-0.627 017
	\bar{X}^2	1.6E-4	1.2E-8	3.4E-17	3.4E-17	3.4E-17	3.4E-17	3.4E-17	3.4E-17
	$\bar{\Delta}^2$	2.3E-3	1.3E-7	(~2 E-14) ^a	(~2 E-14) ^a	(~2 E-14) ^a	(~2 E-14) ^a	(~2 E-14) ^a	(~2 E-14) ^a
C. Cubic	Energy	-0.626 226	-0.627 017	-0.627 017	-0.627 017	-0.627 017	-0.627 017	-0.627 017	-0.627 017
	\bar{X}^2	1.6E-4	4.3E-10	5.3E-21	5.3E-21	5.3E-21	5.3E-21	5.3E-21	5.3E-21
	$\bar{\Delta}^2$	2.3E-3	8.6E-9	(~7 E-25) ^a	(~7 E-25) ^a	(~7 E-25) ^a	(~7 E-25) ^a	(~7 E-25) ^a	(~7 E-25) ^a

^a Estimated.

TABLE II. Convergence of the GVB (3) wave function for the lowest singlet state of FeO₂. The starting vectors were obtained from a partially converged calculation. Add -170.473 086 to the energy quoted to obtain the total energy (with the Fe core electrons replaced by the effective potential).

Iteration	1	2	3	4	5	6	7	8
Method								
A. Uncoupled								
Energy	-0.422 487	-0.538 194	-0.553 252	-0.556 851	-0.557 921	-0.558 362	-0.558 535	-0.558 608
\bar{X}^2	8.9E-2	1.0E-2	2.8E-3	6.8E-4	2.8E-4	9.6E-5	4.4E-5	1.6E-5
$\bar{\Delta}^2$	9.6E-2	3.1E-2	1.1E-2	5.2E-3	2.4E-3	1.0E-3	4.3E-4	1.8E-4
B. Quadratic								
Energy	-0.422 487	-0.554 568	-0.558 635	-0.558 658	-0.558 658			
\bar{X}^2	8.9E-2	2.3E-3	7.4E-6	2.7E-11	2.3E-22			
$\bar{\Delta}^2$	8.6E-2	3.6E-3	3.1E-5	1.0E-10	6.6E-21			
C. Cubic								
Energy	-0.422 487	-0.557 547	-0.558 658	-0.558 658				
\bar{X}^2	8.9E-2	7.0E-4	1.2E-9	7.3E-23				
$\bar{\Delta}^2$	4.8E-2	9.4E-4	1.1E-6	6.3E-21				

(15 occupied orbitals of which six were involved in GVB pairs).¹⁷ In this case the trial functions were obtained from a partially converged calculation.

For NiCH₂ we also used an effective potential¹⁸ to replace the Ni core electrons reducing the system to 18 electrons.¹⁹ A double zeta basis was used (34 basis functions) and we solved for the lowest triplet state HF wave function. We used the results of a converged GVB(2) calculation¹⁵ on the lowest quintet state for the trial function.

One major problem was encountered in the last case (NiCH₂). The trial wave function was found to converge quickly to a stationary point (see Table III, B and C). However, instead of being the lowest minimum, this point was found to be a saddle point. (The first excited state of NiCH₂ is only 0.05 eV above the ground state, and there are a number of excited states within 1.5 eV.)

In order to force convergence to a minimum, one must be able to determine what type of stationary point one is converging toward. The matrix \underline{B} corresponds exactly to the second derivatives of the energy, and thus the eigenvalues of \underline{B} give the principal curvatures of the energy surface [in the $N(N-1)/2$ dimensional rotation space]. Consequently, a necessary condition for convergence to a minimum is that all eigenvalues of the \underline{B} matrix be positive. If any of the eigenvalues of the matrix \underline{B} becomes negative, this indicates that the orbitals are outside the radius of convergence for the lowest minimum. As a result, the expansion of the energy to second or third order about the current set of orbitals is probably insufficient to predict the position of the lowest minimum.

Fortunately, it has been found (empirically) that in such a case it is usually sufficient to move an arbitrary amount (typically between $\frac{1}{6}\pi$ and $\frac{1}{2}\pi$) in the direction of the eigenvector corresponding to the lowest (negative) eigenvalue of the matrix \underline{B} . In this way the program searches for a region of rotation space where the energy has positive curvature, at which point it may then converge cubically (or quadratically). (See Table III.D.) This has worked extremely well in cases to date even when several orbitals have had to be interchanged in order to reach the correct minimum.

The same technique may also be used when solving for excited states in order to prevent convergence to stationary points with incorrect curvature (i.e., the ground state).

V. CONCLUSION

We have shown that by expanding a unitary transformation matrix in terms of completely independent rotation angles one may derive variational equations correct to any order for general open-shell correlated wave functions. Use of these equations leads to extremely rapid convergence of the wave function. In addition, because the derivatives of the energy surface are rigorously evaluated, one may force convergence to a stationary point corresponding to either the ground state or to specified excited states.

ACKNOWLEDGMENT

We thank David H. Munro for helpful discussions.

TABLE III. Convergence of the HF wave function for the lowest triplet state of NiCH_2 . The starting vectors were obtained from the GVB (2) wave function for the lowest quintet state. Add -78.583298 to the energy to obtain the total energy (with the Ni core electrons replaced by the effective potential).

Iteration	1	20	40	60	80	100	120	140	160	174
A. Hamiltonian diagonalization										
Energy	-0.747228	-0.779787	-0.786744	-0.787601	-0.798821	-0.815075	-0.822394	-0.823752	-0.823998	-0.823968
\bar{X}^2	5.8E-3	1.2E-3	5.1E-5	1.2E-4	1.0E-3	8.5E-4	2.0E-4	2.9E-5	3.6E-6	9.9E-7
$\bar{\Delta}^2$										
B. Quadratic										
Energy	-0.747228	-0.605000	-0.028015	-0.001757	-0.003887	-0.006817	-0.007102	-0.007103	-0.007103	
\bar{X}^2	2.1E-4	3.1E-2	1.4E-2	5.6E-3	2.5E-3	3.5E-5	1.4E-7	2.7E-13	1.7E-23	
$\bar{\Delta}^2$	1.3E-0 ^a	2.6E-0 ^a	7.3E-2 ^a	3.1E-1 ^a	2.5E-2 ^a	2.4E-3 ^a	4.4E-6 ^a	2.8E-12 ^a	4.2E-23 ^a	
C. Cubic										
Energy	-0.747228	-0.723710	-0.731686	-0.727862	-0.727856	-0.727850				
\bar{X}^2	2.1E-4	2.8E-3	2.3E-4	5.3E-7	7.9E-14	1.4E-30				
$\bar{\Delta}^2$	6.9E-2 ^a	3.9E-2 ^a	5.9E-2 ^a	9.5E-5 ^a	1.3E-12 ^a	3.3E-28 ^a				
D. Quadratic, with convergence forced to region of positive curvature										
Energy	-0.747228	-0.769040	-0.769263	-0.793843	-0.798544	-0.807581	-0.823551	-0.823974	-0.823975	-0.823976
\bar{X}^2	2.1E-4	3.0E-4	2.4E-3	1.6E-3	1.6E-3	2.5E-3	4.5E-5	1.3E-7	3.0E-12	1.7E-21
$\bar{\Delta}^2$	5.0E-1 ^b	5.0E-1	5.0E-1 ^b	5.0E-1 ^b	5.0E-1 ^b	5.0E-2	2.4E-3	1.2E-5	3.4E-10	(-1 E-19) ^c

^a The B matrix contains at least one negative eigenvalue.

^b Negative curvature was found in the B matrix and handled by the procedure given in Sec. IV C.

^c Estimated.

APPENDIX A: EXPANDED VARIATIONAL EQUATIONS

The general variational equations of (38) and (35) contain permutation operators. This appendix lists the expanded form of these equations, with the permutation operators applied. The variational equations have the form

$$0 = X_{\mu\nu} + \Delta_{mn}(B_{\mu\nu} + C_{\mu\nu mn pq} \Delta_{pq})$$

with $\mu > \nu$, $m > n$, but with the sum over p and q remaining unrestricted.

1. General energy expression

From (27) we have

$$X_{\mu\nu} = H_{\nu j}^{\mu j} - H_{\mu j}^{\nu j}, \quad (\text{A1})$$

where $\mu > \nu$;

$$B_{\mu\nu mn} = \{ [H_{\nu n}^{\mu m} + 2V_{\nu jkn}^{\mu jkm} + \frac{1}{2}\delta_{mn}(H_{nj}^{\mu j} + H_{\mu j}^{\nu j})] + [H_{\mu m}^{\nu n} + 2V_{\mu jkm}^{\nu jkn} + \frac{1}{2}\delta_{mn}(H_{mj}^{\nu j} + H_{\nu j}^{\mu j})] \\ - [H_{\mu n}^{\nu m} + 2V_{\mu jkn}^{\nu jkm} + \frac{1}{2}\delta_{mn}(H_{nj}^{\nu j} + H_{\mu j}^{\nu j})] - [H_{\nu m}^{\mu n} + 2V_{\nu jkm}^{\mu jkn} + \frac{1}{2}\delta_{mn}(H_{mj}^{\mu j} + H_{\nu j}^{\mu j})] \}, \quad (\text{A2})$$

where $\mu > \nu$, $m > n$;

$$C_{\mu\nu mn pq} \Delta_{pq} = \{ (V_{\nu nqj}^{\mu mpj} + V_{\mu mqj}^{\nu npj} - V_{\mu nqj}^{\nu mpj} - V_{\nu mqj}^{\mu npj}) \Delta_{pq} + (V_{\nu qmj}^{\mu pmj} + V_{\mu qmj}^{\nu pmj} - V_{\mu qmj}^{\nu pmj} - V_{\nu qmj}^{\mu pmj}) \Delta_{pq} + (V_{nq\mu j}^{\mu p\mu j} + V_{mq\mu j}^{\nu p\mu j} - V_{nq\mu j}^{\nu p\mu j} - V_{mq\mu j}^{\mu p\mu j}) \Delta_{pq} \} \\ + \delta_{m\nu} \{ [\frac{1}{4}(H_{nq}^{\mu p} + H_{\mu q}^{\nu p}) + \frac{1}{2}(V_{njka}^{\mu jkp} + V_{\mu jka}^{\nu jkp})] \Delta_{pq} + \frac{1}{12} [(H_{qj}^{\mu j} - H_{\mu j}^{\nu j}) \Delta_{nq} + (H_{qj}^{\nu j} - H_{\mu j}^{\mu j}) \Delta_{\mu q}] \} \\ + \delta_{\mu n} \{ [\frac{1}{4}(H_{mq}^{\nu p} + H_{\nu q}^{\mu p}) + \frac{1}{2}(V_{mjka}^{\nu jkp} + V_{\nu jka}^{\mu jkp})] \Delta_{pq} + \frac{1}{12} [(H_{qj}^{\nu j} - H_{\nu j}^{\mu j}) \Delta_{mq} + (H_{qj}^{\mu j} - H_{\nu j}^{\nu j}) \Delta_{\nu q}] \} \\ - \delta_{\mu m} \{ [\frac{1}{4}(H_{nq}^{\nu p} + H_{\nu q}^{\mu p}) + \frac{1}{2}(V_{njka}^{\nu jkp} + V_{\nu jka}^{\mu jkp})] \Delta_{pq} + \frac{1}{12} [(H_{qj}^{\nu j} - H_{\nu j}^{\mu j}) \Delta_{nq} + (H_{qj}^{\mu j} - H_{\nu j}^{\nu j}) \Delta_{\nu q}] \} \\ - \delta_{\nu n} \{ [\frac{1}{4}(H_{mq}^{\mu p} + H_{\mu q}^{\nu p}) + \frac{1}{2}(V_{mjka}^{\mu jkp} + V_{\mu jka}^{\nu jkp})] \Delta_{pq} + \frac{1}{12} [(H_{qj}^{\mu j} - H_{\mu j}^{\nu j}) \Delta_{mq} + (H_{qj}^{\nu j} - H_{\mu j}^{\mu j}) \Delta_{\mu q}] \} \\ + [\frac{1}{4}(H_{\mu n}^{\nu q} - H_{\nu n}^{\mu q} + H_{\mu q}^{\nu n} - H_{\nu q}^{\mu n}) + \frac{1}{2}(V_{\mu jka}^{\nu jkp} - V_{\nu jka}^{\mu jkp} + V_{\mu jka}^{\nu jkp} - V_{\nu jka}^{\mu jkp})] \Delta_{mq} \\ + [\frac{1}{4}(H_{\nu m}^{\mu q} - H_{\mu m}^{\nu q} + H_{\nu q}^{\mu m} - H_{\mu q}^{\nu m}) + \frac{1}{2}(V_{\nu jka}^{\mu jkp} - V_{\mu jka}^{\nu jkp} + V_{\nu jka}^{\mu jkp} - V_{\mu jka}^{\nu jkp})] \Delta_{nq} \\ + [\frac{1}{4}(H_{mq}^{\nu n} - H_{nq}^{\mu \nu} + H_{\nu m}^{\nu q} - H_{\nu m}^{\mu q}) + \frac{1}{2}(V_{mjka}^{\nu jkp} - V_{mjka}^{\mu jkp} + V_{mjka}^{\nu jkp} - V_{mjka}^{\mu jkp})] \Delta_{\mu q} \\ + [\frac{1}{4}(H_{nq}^{\mu \nu} - H_{nq}^{\nu \mu} + H_{\nu m}^{\mu q} - H_{\nu m}^{\nu q}) + \frac{1}{2}(V_{njka}^{\mu jkp} - V_{njka}^{\nu jkp} + V_{njka}^{\mu jkp} - V_{njka}^{\nu jkp})] \Delta_{\nu q} \\ + \frac{1}{12} [(H_{\mu j}^{\nu j} - H_{\nu j}^{\mu j}) \Delta_{m\nu} + (H_{m j}^{\nu j} - H_{\nu j}^{\mu j}) \Delta_{\mu n} - (H_{\mu j}^{\mu j} - H_{\nu j}^{\nu j}) \Delta_{\nu n} - (H_{\nu j}^{\nu j} - H_{\mu j}^{\mu j}) \Delta_{\mu m}], \quad (\text{A3})$$

where $\mu > \nu$, $m > n$, and the sums over p and q are not restricted.

2. Restricted energy expression

From (35) we have

$$X_{\mu\nu} = H_{\mu\nu}^{\mu\nu} - H_{\mu\nu}^{\nu\nu}, \quad (\text{A4})$$

where $\mu > \nu$;

$$B_{\mu\nu mn} = \{ \delta_{\mu n} [H_{\nu n}^{\mu\nu} - \frac{1}{2}(H_{\nu n}^{\mu\nu} + H_{\nu n}^{\nu\nu})] + 2g_{\mu\nu mn}^{\mu\nu} + \delta_{\nu n} [H_{\mu m}^{\nu\nu} - \frac{1}{2}(H_{\mu m}^{\mu\nu} + H_{\mu m}^{\nu\nu})] + 2g_{\mu\nu mn}^{\nu\nu} \\ - \delta_{m\nu} [H_{\mu n}^{\nu\nu} - \frac{1}{2}(H_{\mu n}^{\mu\nu} + H_{\mu n}^{\nu\nu})] - 2g_{\mu\nu mn}^{\mu\nu} - \delta_{\mu n} [H_{\nu m}^{\mu\nu} - \frac{1}{2}(H_{\nu m}^{\mu\nu} + H_{\nu m}^{\nu\nu})] - 2g_{\mu\nu mn}^{\nu\nu} \}, \quad (\text{A5})$$

where $\mu > \nu$, $m > n$;

$$\begin{aligned}
C_{\mu\nu m n p q} \Delta_{p q} = & \{ [g_{\mu\nu n q}^{\mu m} - g_{\mu\nu n q}^{\nu m} + \frac{1}{2}(g_{\mu\nu n q}^{\nu n} - g_{\mu\nu n q}^{\mu n} + g_{\mu\nu n q}^{\nu q} - g_{\mu\nu n q}^{\mu q})] \Delta_{m q} \\
& + [g_{m n \nu q}^{\mu m} - g_{m n \nu q}^{\nu m} + \frac{1}{2}(g_{m n \nu q}^{\nu n} - g_{m n \nu q}^{\mu n} + g_{m n \nu q}^{\nu q} - g_{m n \nu q}^{\mu q})] \Delta_{\mu q} \\
& - [g_{\mu\nu m q}^{\mu m} - g_{\mu\nu m q}^{\nu m} + \frac{1}{2}(g_{\mu\nu m q}^{\nu n} - g_{\mu\nu m q}^{\mu n} + g_{\mu\nu m q}^{\nu q} - g_{\mu\nu m q}^{\mu q})] \Delta_{n q} \\
& - [g_{m n \mu q}^{\mu m} - g_{m n \mu q}^{\nu m} + \frac{1}{2}(g_{m n \mu q}^{\nu n} - g_{m n \mu q}^{\mu n} + g_{m n \mu q}^{\nu q} - g_{m n \mu q}^{\mu q})] \Delta_{\nu q} \\
& + [\frac{1}{4}(H_{\nu n}^m - H_{\nu n}^\mu) + \frac{1}{12}(H_{\nu n}^\nu - H_{\nu n}^n)] \Delta_{\mu m} + [\frac{1}{4}(H_{\mu m}^n - H_{\mu m}^\nu) + \frac{1}{12}(H_{\mu m}^\mu - H_{\mu m}^m)] \Delta_{\nu n} \\
& + [\frac{1}{4}(H_{\mu n}^\nu - H_{\mu n}^m) + \frac{1}{12}(H_{\mu n}^n - H_{\mu n}^\mu)] \Delta_{m \nu} + [\frac{1}{4}(H_{\nu m}^n - H_{\nu m}^\mu) + \frac{1}{12}(H_{\nu m}^\nu - H_{\nu m}^m)] \Delta_{\mu n} \\
& + \delta_{\mu m} [(\frac{1}{4}(H_{q n}^\mu - H_{q n}^\nu) + \frac{1}{12}(H_{q n}^q - H_{q n}^n)) \Delta_{\nu q} + (\frac{1}{4}(H_{q \nu}^m - H_{q \nu}^n) + \frac{1}{12}(H_{q \nu}^q - H_{q \nu}^\nu)) \Delta_{n q} + (g_{p q n \nu}^{\mu m} - \frac{1}{2}(g_{p q n \nu}^{\mu n} + g_{p q n \nu}^{\nu \mu})) \Delta_{p q}] \\
& + \delta_{\nu n} [(\frac{1}{4}(H_{q m}^\nu - H_{q m}^\mu) + \frac{1}{12}(H_{q m}^q - H_{q m}^m)) \Delta_{\mu q} + (\frac{1}{4}(H_{q \mu}^n - H_{q \mu}^m) + \frac{1}{12}(H_{q \mu}^q - H_{q \mu}^n)) \Delta_{m q} + (g_{p q m \mu}^{\nu m} - \frac{1}{2}(g_{p q m \mu}^{\nu n} + g_{p q m \mu}^{\mu \nu})) \Delta_{p q}] \\
& - \delta_{m \nu} [(\frac{1}{4}(H_{q n}^\nu - H_{q n}^\mu) + \frac{1}{12}(H_{q n}^q - H_{q n}^n)) \Delta_{\mu q} + (\frac{1}{4}(H_{q \mu}^m - H_{q \mu}^n) + \frac{1}{12}(H_{q \mu}^q - H_{q \mu}^m)) \Delta_{n q} + (g_{p q n \mu}^{\mu m} - \frac{1}{2}(g_{p q n \mu}^{\mu n} + g_{p q n \mu}^{\nu \mu})) \Delta_{p q}] \\
& - \delta_{\mu n} [(\frac{1}{4}(H_{q m}^\mu - H_{q m}^\nu) + \frac{1}{12}(H_{q m}^q - H_{q m}^m)) \Delta_{\nu q} + (\frac{1}{4}(H_{q \nu}^n - H_{q \nu}^m) + \frac{1}{12}(H_{q \nu}^q - H_{q \nu}^\nu)) \Delta_{m q} + (g_{p q m \nu}^{\nu m} - \frac{1}{2}(g_{p q m \nu}^{\nu n} + g_{p q m \nu}^{\mu \nu})) \Delta_{p q} \},
\end{aligned}
\tag{A6}$$

where $\mu > \nu$, $m > \mu$ and the sums over p and q are not restricted.

*Partially supported by grants MPS74-05132 and DMR74-04965 from the National Science Foundation.

†National Science Foundation Undergraduate Research Participant, summer 1975.

¹W. J. Hunt, P. J. Hay, and W. A. Goddard III, *J. Chem. Phys.* **57**, 738 (1972).

²C. C. J. Roothaan, *Rev. Mod. Phys.* **32**, 179 (1960); R. K. Nesbet, *Proc. Roy. Soc. Lond. A* **230**, 312 (1955); F. W. Birss and S. Fraga, *J. Chem. Phys.* **38**, 2552 (1963).

³W. A. Goddard III, T. H. Dunning, Jr., and W. J. Hunt, *Chem. Phys. Lett.* **4**, 231 (1969).

⁴W. J. Hunt, W. A. Goddard III, and T. H. Dunning, Jr., *Chem. Phys. Lett.* **6**, 147 (1970).

⁵F. W. Bobrowicz and W. A. Goddard III, in *Electronic Structure, Vol. II of Modern Theoretical Chemistry*, edited by H. F. Schaefer III (Plenum, New York, to be published).

⁶F. W. Bobrowicz, Ph.D. thesis (California Institute of Technology, 1974) (unpublished).

⁷If complex orbitals are allowed, $\underline{\Delta}$ must be an arbitrary anti-Hermitian matrix.

⁸Because $\underline{T} \underline{T}^\dagger = e^{\underline{\Delta}} e^{\underline{\Delta}^\dagger} = e^{\underline{\Delta} + \underline{\Delta}^\dagger} = e^{\underline{\Delta} - \underline{\Delta}} = e^{\underline{0}} = \underline{1}$, we find that \underline{T} is unitary (as required).

⁹Without loss of generality, we may assume that c_{ijkl} has the same symmetry as $(ij|kl)$. That is,

$$c_{ijkl} = c_{jkl i} = c_{ijlk} = c_{klij} = c_{klji} = c_{ikij} = c_{ikji}.$$

¹⁰The permutation operators for X_{mn} , B_{mnpq} , and C_{mnpqrs} are $[n]_1$, $[n]_2$, and $[n]_3$, respectively. The general form is $O_{ff}^{[2n]}$ [W. A. Goddard III, *Phys. Rev.* **157**, 73 (1967)].

¹¹ $X_{\mu\nu}$ and $B_{\mu\nu m n}$ had been derived previously in essential-

ly the same form by Hunt *et al.*⁴ Their B matrix is slightly different due to differences of $O(\Delta^2)$ in the definition of \underline{T} in terms of $\underline{\Delta}$.

^{12(a)}The GVBONE program¹ uses the first-order form of (37a),

$$0 = X_{21} + \Delta_{21} B_{221} \tag{37b}$$

and the form of X and B given in (37e) and (37f).

^{12(b)}Our uncoupled approach differs from that in GVBONE in two respects: (i) we use (37b) also for the virtual orbitals (rather than OCBSE) and (ii) after obtaining the Δ_{ij} from (37b) we insert in (13) to obtain the transformation matrix, whereas the usual approach is to obtain new orbitals using $T = 1 + \underline{\Delta}$ and then reorthogonalizing.

¹³Throughout Sec. IV, vectors (such as \bar{X} or $\bar{\Delta}$) and tensors (such as \underline{B} or \underline{C}) are assumed to reside in the $N(N-1)/2$ dimensional rotation space. However, as noted in Sec. II, $\bar{\Delta}$ is equivalent to an N -dimensional antisymmetric matrix $\underline{\Delta}$ which is used to construct the transformation \underline{T} .

¹⁴T. H. Dunning, *J. Chem. Phys.* **53**, 2823 (1970).

¹⁵A GVB (p) wave function differs from the HF wave function in that p doubly occupied orbitals (of HF) are described as correlated pairs.¹

¹⁶C. F. Melius, B. D. Olafson, and W. A. Goddard III, *Chem. Phys. Lett.* **28**, 457 (1974).

¹⁷R. Bair, W. A. Goddard III, and B. D. Olafson (unpublished).

¹⁸M. J. Sollenberger, W. A. Goddard III, and C. F. Melius (unpublished).

¹⁹A. Rappé and W. A. Goddard III (unpublished).